View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Online Research Database In Technology

DTU Library

Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans

Moreno-Mayar, J Víctor; Potter, Ben A; Vinner, Lasse; Steinrücken, Matthias; Rasmussen, Simon; Terhorst, Jonathan; Kamm, John A; Albrechtsen, Anders; Malaspinas, Anna-Sapfo; Sikora, Martin

Published in:
Nature

Link to article, DOI:
10.1038/nature25173

Publication date:
2017

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
Moreno-Mayar, J. V., Potter, B. A., Vinner, L., Steinrücken, M., Rasmussen, S., Terhorst, J., ... Willerslev, E. (2017). Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans. *Nature*, *553*(7687), 203-207. https://doi.org/10.1038/nature25173

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LIVERPOOL JOHN MOORES UNIVERSITY

LJMU Research Online

Moreno-Mayar, JV, Potter, BA, Vinner, L, Steinrücken, M, Rasmussen, S, Terhorst, J, Kamm, JA, Albrechtsen, A, Malaspinas, A-S, Sikora, M, Reuther, JD, Irish, JD, Malhi, RS, Orlando, L, Song, YA, Nielsen, R, Meltzer, DJ and Willerslev, E

Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans

http://researchonline.ljmu.ac.uk/7887/

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Moreno-Mayar, JV, Potter, BA, Vinner, L, Steinrücken, M, Rasmussen, S, Terhorst, J, Kamm, JA, Albrechtsen, A, Malaspinas, A-S, Sikora, M, Reuther, JD, Irish, JD, Malhi, RS, Orlando, L, Song, YA, Nielsen, R, Meltzer, DJ and Willerslev, E (2018) Terminal Pleistocene Alaskan genome reveals first

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/

http://researchonline.ljmu.ac.uk/

Terminal Pleistocene Alaskan genome reveals first founding population of Native Americans

J. Víctor Moreno-Mayar [1,*], Ben A. Potter [2,*], Lasse Vinner [1,*] Matthias Steinrücken [3,4,5], Simon Rasmussen [6], Jonathan Terhorst [4], John A. Kamm[4,7], Anders Albrechtsen [8], Anna-Sapfo Malaspinas [1,9,10], Martin Sikora [1], Joshua D. Reuther[2], Joel D. Irish[11], Ripan S. Malhi [12,13], Ludovic Orlando [1], Yun S. Song [3,4,14,15], Rasmus Nielsen [1,4,14], David J. Meltzer [1,16] and Eske Willerslev [1,7,17,].**

1. Centre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark.
2. Department of Anthropology, University of Alaska, Fairbanks, AK 99775.
3. Computer Science Division, University of California, Berkeley, CA 94720, USA.
4. Department of Statistics, University of California, Berkeley, CA 94720, USA.
5. Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, MA 01003, USA.
6. Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, 2800 Kongens Lyngby, Denmark.
7. Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge, CB10 1SA, UK.
8. The Bioinformatics Centre, Department of Biology, University of Copenhagen, 2200 Copenhagen, Denmark.
9. Institute of Ecology and Evolution, University of Bern, CH-3012 Bern, Switzerland.
10. Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland.
11. Research Centre in Evolutionary Anthropology and Palaeoecology, Liverpool John Moores University, Liverpool L3 3AF, UK.
12. Department of Anthropology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
13. Carle R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.
14. Department of Integrative Biology, University of California, Berkeley, CA 94720, USA.
15. Department of Mathematics and Department of Biology, University of Pennsylvania, PA 19104.
16. Department of Anthropology, Southern Methodist University, Dallas, TX 75275, USA.
17. Department of Zoology, University of Cambridge, Downing St, Cambridge CB2 3EJ, UK

* These authors contributed equally to this work.
** Corresponding author: ewillerslev@snm.ku.dk

Despite broad agreement that the Americas were initially populated via Beringia, when and how this happened is debated [1–5]. Key to this debate are human remains from Late Pleistocene Alaska. The first and only such remains were recovered at Upward Sun River (USR), and date to ~11.5 kya [6,7]. We sequenced the USR1 genome to an average coverage of ~17X. We find USR1 is most closely related to Native Americans, but falls basal to all previously sequenced contemporary and ancient Native Americans [1,8,9]. As such, USR1 represents a distinct Ancient Beringian (AB)

population. Using demographic modelling we infer the AB population and ancestors of other Native Americans descend from a single founding population that initially split from East Asians ~36 ± 1.5 kya, with gene flow persisting until ~25 ± 1.1 kya. Gene flow from ancient north Eurasians into all Native Americans took place 25-20 kya, with AB branching off ~22-18.1 kya. Our findings support long-term genetic structure in ancestral Native Americans, consistent with the Beringian Standstill Model [10]. We find that the basal Northern (NNA) and Southern (SNA) branches, to which all other Native Americans belong, diverged ~17.5-14.6 kya, likely south of the North American ice sheets. After 11.5 kya, some NNA populations received gene flow from a Siberian population most closely related to Koryaks, but not Paleoeskimos[1], Inuit or Kets [11], and that Native American gene flow into Inuit was via NNA and not SNA groups[1]. Our findings further suggest the far northern North American presence of NNA is from a back migration that replaced or absorbed the initial AB founding population.

The peopling of the Americas, and particularly the population history of Beringia, the land bridge that connected far northeast Asia to northwestern North America during the Pleistocene, remains unresolved [2,3]. Humans were present in the Americas south of the continental ice sheets by ~14.6 kya [12], indicating they traversed Beringia earlier, possibly around the Last Glacial Maximum (LGM). Then, the region was marked by harsh climates and glacial barriers [5], which may have led to the isolation of populations for extended periods, and at times complicated dispersal across the region [13]. Still controversial are questions of whether and how long Native American ancestors were isolated from Asian groups in Beringia prior to entering the Americas [2,10,14]; if one or more early migrations gave rise to the founding population of Native Americans [1–4,8,15] (it is commonly agreed Paleoeskimos and Inuit represent separate and later migrations [1,16,17]); and, when and where the basal split between SNA and NNA occurred. Unresolved too is whether the genetic affinity between some SNA groups and indigenous Australasians [2,3], reflects migration by non-Native Americans [3,4,15], early population structure within the first Americans [3], or later gene flow [2]. Key to resolving these uncertainties is a better understanding of the population history of Beringia, the entryway for the Pleistocene peopling of the Americas.

Genomic insight into that population history has now become available with the recently recovered infant remains (USR1 and USR2) from the Upward Sun River site, Alaska (eastern Beringia), dated to ~11.5 kya [7,18]. Mitochondrial DNA sequences (haplogroups C1 and B2, respectively) were previously acquired from these individuals [7,18] (SI 1,4.5). We have since obtained whole-genome sequence data, which provides a broader opportunity to investigate the number, source(s) and structure of the initial founding population(s), and the timing and location of their subsequent divergence. We sequenced the genome of USR1 to an average depth of ~17X, based on eight sequencing libraries from USER-treated extracts previously confirmed to contain DNA fragments with characteristic ancient DNA misincorporation patterns (SI 2-4). We estimated modern human contamination at ~0.14% based on the nuclear genome and ~0.15% based on

mtDNA (SI 4.). As expected, the error rate in the USER-treated sequencing data was low (0.09% errors per-base), and comparable to other high-coverage contemporary genomes, based on called genotypes (SI 4). While USR2 [7] did not show sufficient endogenous DNA for high-coverage genome sequencing, we found both individuals were close relatives (SI 5), equally related to worldwide present-day populations (Figure S4g).

We assessed the genetic relationship between USR1, a set of ancient genomes [2,8,9,15,17], and a panel of 167 worldwide populations genotyped for 199,285 SNPs [1,2,19] (SI 6), using outgroup *f3* statistics [20], model-based clustering [21,22] and multidimensional scaling (MDS) [23] (SI 7-9). Outgroup *f3* statistics of the form *f3(Yoruba; X, USR1)* revealed that USR1 is more closely related to present-day Native Americans than to any other tested population, followed by Siberian and East Asian populations [1,2] (Figure 1a). Pairwise comparisons of the *f3*-statistics for USR1 and a set of ancient and contemporary Native American genomes [2,8,15] (SI 6) showed that all are similarly related to Old World populations, though other Native American genomes (Aymara [2], Athabascan1 [16], 939 [2], Anzick1 [8] and Kennewick [15]) have a higher affinity for contemporary Native Americans than USR1 does (SI 9). MDS and *ADMIXTURE* analysis showed that the USR1 genome did not cluster with any specific Native American group (Figures 1d, S3b). These results imply that USR1 belonged to a previously unknown Native American population not represented in the reference dataset, herein identified as Ancient Beringians (SI 8.3).

To investigate if USR1 derived from the same source population that gave rise to contemporary Native Americans, we computed 11,322 allele frequency based-*D*-statistics [1,20] of the form *D(Native American, USR1; Siberian1/Han, Siberian2/Han)* (SI 10.4). The resulting *Z*-score distribution corresponds qualitatively to the expected normal distribution under the null hypothesis that USR1 forms a clade with Native Americans to the exclusion of Siberians and East Asians – except for a set of Eskimo-Aleut, Athabascan and Northern Amerind-speaking populations for which recent Asian gene flow has been previously documented (Figures 1c, S5a, S6) [1,2,15,19]. Additionally, we found that present-day Native Americans and USR1 yield similar results for *D(Native American/USR1, Han; Mal'Ta, Yoruba)*, suggesting they are equally related to the ancient north Eurasian population represented by the 24 kya Mal'ta individual [9] (SI 10.5). These results confirm that USR1 and present-day Native Americans derived from the same ancestral source, which carried a mixture of East Asian and Mal'ta-related ancestry. We infer that descendants of this source represent the basal group that first migrated into the Americas.

To explore the relationship between USR1 and present-day Native Americans, we computed allele frequency-based and genome-wide *D*-statistics of the form *D(Native American, Aymara; USR1, Yoruba)*. We could not reject the null hypothesis that USR1 is an outgroup to any pair of Native Americans, with the exception of a set of populations bearing recent Asian gene flow [1,2,15,19] (Figures 1b, S7). We confirmed the phylogenetic placement of USR1 at a basal position in the Native American clade using TreeMix [24] and two methods to estimate average genomic divergence and genetic drift, respectively (SI 14-16). These results support the branching of USR1 within the Native American

clade, but being equidistant to NNA and SNA. Below we discuss the potential geographic locations of the USR1-NNA+SNA and the NNA-SNA splits (Figure 2) based on the genetic results, the glacial geography of terminal Pleistocene North America [25,26] and the extant archaeological evidence (also SI 20).

Recent detection of an Australasian-derived genetic signature in some Native American groups [2,3] led us to explore whether USR1 bears that signal (SI 10.7, 11-13). Using frequency-based and 'enhanced' D-statistics, we found no support for USR1 being closer to Papuans (a proxy for Australasians) than other Native Americans.

We leveraged the position of USR1 on the Native American branch prior to the NNA-SNA split to re-assess the origins of Athabascan and Eskimo populations by fitting admixture graphs. We considered a whole-genome dataset including Siberian, East Asian, Native American and Eskimo groups, as well as Mal'ta (SI 17). The heuristic approach in TreeMix [24] showed that the best proxies for the Asian component in Athabascans and Greenlandic Inuit are Koryaks and the Saqqaq individual, respectively. We then followed an incremental approach for fitting an *f*-statistic-based admixture graph [20], including the Kets, previously suggested to share a linguistic and perhaps a genetic link with Athabascans [11,27]. This approach recapitulated the TreeMix results , and yielded a model in which both Athabascans and Greenlandic Inuit derive from the NNA branch. However, the Asian ancestry in Athabascans is most closely related to the Asian component in Koryaks, while the Saqqaq genome is the best proxy for the Siberian component in the Greenlandic Inuit (Figure 3). We infer the latter is a consequence of Palaeo- and Neo-Eskimos having been derived from a similar Siberian population [1,16]. This model appears to be a good fit to the data, as the observed *f*-statistic that deviated the most from the model prediction yielded Z=3.27. In SI 17.3 we tested the robustness of this model and predictions by computing individual D statistics, and re-fitting the model using alternative datasets.

Lastly, we inferred the demographic history of USR1 with respect to Native Americans, Siberians and East Asians, using two independent methods: *diCal2* [28] and *momi2* [29] (SI 18-19). *diCal2* results indicate that the founding population of USR1, Native Americans, and Siberians had a very weak structure from ~36 kya up to ~24.5 kya (Table S7), when the ancestors of USR1 and Native Americans began to diverge substantially from Siberians. USR1 diverged from other Native Americans around 20.9 kya, with a period of ensuing moderate gene flow between them (Table S6 and S7), as indicated by a simulation study that showed a significant increase in likelihood when comparing a 'clean split' model to an 'isolation with migration' model (SI 18.4). Using *momi2* and *SMC++* we estimated a backbone demography where Karitiana and Athabascans split at ~15.7 kya, while their ancestral population split from Koryaks ~23.3 kya (Figure 4). With *momi2,* we inferred the most likely branch (the population immediately ancestral to NNA+SNA) and time (~21 kya) for the USR1 population to join the backbone demography, while allowing for possible gene flow between USR and other populations (SI 19, Figure 4b), results consistent with [14] and the *diCal2* inference.

These new findings, along with existing data, allow us to place Ancient Beringians (AB) within the broader context of the Pleistocene peopling of the Americas. The Native American founding population (comprised of both AB and NNA+SNA) began to diverge from ancestral Asians as early as ~36 kya, likely in northeast Asia, as there is no evidence of people in Beringia or northwest North America at this period. A high level of gene flow was maintained between them and other Asians until as late as ~25 kya [2,14]. The subsequent isolation of the Native American founding population ~24 kya roughly corresponds with a decline in archaeological evidence for a human presence in Siberia [30]. Both changes may result from the same underlying cause: the onset of harsh LGM climatic conditions [2]. These findings, coupled with a divergence date of ~20.9 kya between USR1 and Native Americans, are in agreement with the Beringian Standstill Model [10] (SI 21). The common ancestor of NNA+SNA and AB began to diverge ~20.9 kya, after which gene flow ensued, although whether it was with NNA+SNA, or the already differentiated NNA and SNA branches, cannot be determined owing to shallow divergence times among the groups.

These findings allow us to consider possible scenarios regarding where ancient Native American populations diverged (SI 20-21, Figure 2). Scenarios C-E require extended periods of strong population structure marking AB, NNA, and SNA as separate groups, for which we do not see compelling genetic evidence; hence these can be rejected. Scenarios A and B are compatible with our evidence of continuous gene flow among these groups, but differ as to the location of the AB versus NNA+SNA split at 20.9 kya, whether in northeast Asia (Scenario A) or eastern Beringia (Scenario B). Each has strengths and weaknesses relative to genetic and archaeological evidence: Scenario A best fits the archaeological and paleoecological evidence, as the earliest securely dated sites in Beringia are no older than ~15-14 kya, and the LGM cold period is unlikely to be associated with northward expanding populations [30]. Scenario B is genetically most parsimonious, given evidence of continuous gene flow between the AB and NNA+SNA, suggesting their geographical proximity 20.9-11.5 kya, and that all three were isolated from Asian/Siberian groups after ~24 kya and form a clade.

Scenarios A and B are both consistent with the NNA-SNA split at ~15 kya [2] having occurred in a region south of eastern Beringia. The ice sheets were then still a significant barrier to movement that would have helped maintain separation from the AB population. While members of the SNA branch have not been documented in regions that were once north of the glacial ice [1,19], NNA groups (including Athabascan-speakers) are present in Alaska today; thus, the latter are likely descendants of a population that moved north sometime after 11.5 kya [26].

The USR1 results provide the first direct genomic evidence that all Native Americans can be traced back to the same source population from a single Late Pleistocene founding event. Descendants of that population were present in eastern Beringia until at least 11.5 kya. By then, however, a separate branch of Native Americans had already established

itself in unglaciated North America, and diverged into the two basal groups that ultimately became the ancestors of most of the indigenous populations of the Americas.

Data availability

Sequence data was deposited in the ENA under accession: PRJEB20398.

Acknowledgements

The Upward Sun River excavations and analysis were conducted under a Memorandum of Agreement (MOA) signed by the State of Alaska, the National Science Foundation, the Healy Lake Tribal Council, and the Tanana Chiefs Conference. We appreciate the cooperation of all parties. We thank Morten Allentoft, Shyam Gopalakrishnan, Thorfinn Korneliussen, Pablo Librado, Jazmín Ramos-Madrigal, Gabriel Renaud and Filipe Vieira for discussions. We thank the Danish National High-throughput Sequencing Centre for assistance in data generation. GeoGenetics members were supported by the Lundbeck Foundation and the Danish National Research Foundation (DNRF94) and KU2016. J.V.M.-M. was supported by Conacyt (Mexico). Samples were recovered during excavations by B.A.P. supported by NSF Grants 1138811 and 1223119. Research supported in part by NIH grant R01-GM094402 (M.S., J.T., J.A.K., and Y.S.S.) and a Packard Fellowship for Science and Engineering (Y.S.S.). D.J.M. is supported by the Quest Archaeological Research Fund. A.-S.M. is supported by the Swiss National Science Foundation and the ERC.

Author Contributions

Project conceived by E.W. and B.A.P., and headed by E.W. and J.V.M.-M. L.V. processed ancient DNA. J.V.M.-M. and S.R. assembled datasets. J.V.M.-M., M.S., J.T., J.A.K. and A.A. analysed genetic data. B.A.P. led the USR field investigation, and B.A.P. and D.J.M. provided anthropological contextualization. B.A.P., J.D.R., and J.D.I. conducted archaeological and bioanthropological work. R.N., Y.S.S., M.Si., A.-S.M., and L.O. supervised bioinformatic and statistical analyses. B.A.P. engaged with indigenous communities. J.V.M.-M., B.A.P., D.J.M and E.W. wrote the manuscript with input from L.V., A.-S.M., M.Si., R.S.M., L.O., Y.S.S, R.N. and remaining authors.

References

1. Reich, D. *et al.* Reconstructing Native American population history. *Nature* **488,** 370–374 (2012).
2. Raghavan, M. *et al.* Genomic evidence for the Pleistocene and recent population history of Native Americans. *Science* **349,** aab3884–aab3884 (2015).
3. Skoglund, P. *et al.* Genetic evidence for two founding populations of the Americas. *Nature* (2015). doi:10.1038/nature14895
4. von Cramon-Taubadel, N., Strauss, A. & Hubbe, M. Evolutionary population history of early Paleoamerican cranial morphology. *Sci. Adv.* **3,** e1602289 (2017).
5. Hoffecker, J. F., Elias, S. A., O'Rourke, D. H., Scott, G. R. & Bigelow, N. H.

Beringia and the global dispersal of modern humans: Beringia and the Global Dispersal of Modern Humans. *Evol. Anthropol. Issues News Rev.* **25,** 64–78 (2016).
6. Potter, B. A., Irish, J. D., Reuther, J. D., Gelvin-Reymiller, C. & Holliday, V. T. A Terminal Pleistocene Child Cremation and Residential Structure from Eastern Beringia. *Science* **331,** 1058–1062 (2011).
7. Potter, B. A., Irish, J. D., Reuther, J. D. & McKinney, H. J. New insights into Eastern Beringian mortuary behavior: A terminal Pleistocene double infant burial at Upward Sun River. *Proc. Natl. Acad. Sci.* **111,** 17060–17065 (2014).
8. Rasmussen, M. *et al.* The genome of a Late Pleistocene human from a Clovis burial site in western Montana. *Nature* **506,** 225–229 (2014).
9. Raghavan, M. *et al.* Upper Palaeolithic Siberian genome reveals dual ancestry of Native Americans. *Nature* **505,** 87–91 (2013).
10. Tamm, E. *et al.* Beringian Standstill and Spread of Native American Founders. *PLoS ONE* **2,** e829 (2007).
11. Flegontov, P. *et al. Na-Dene populations descend from the Paleo-Eskimo migration into America*. (2016).
12. Dillehay, T. D. *et al.* Monte Verde: seaweed, food, medicine, and the peopling of South America. *Science* **320,** 784–786 (2008).
13. Goebel, T. & Potter, B. A. First Traces: Late Pleistocene Human Settlement of the Arctic. in *The Oxford handbook of the prehistoric Arctic* 223–252 (Oxford University Press, 2016).
14. Llamas, B. *et al.* Ancient mitochondrial DNA provides high-resolution time scale of the peopling of the Americas. *Sci. Adv.* **2,** e1501385–e1501385 (2016).
15. Rasmussen, M. *et al.* The ancestry and affiliations of Kennewick Man. *Nature* (2015). doi:10.1038/nature14625
16. Raghavan, M. *et al.* The genetic prehistory of the New World Arctic. *Science* **345,** 1255832–1255832 (2014).
17. Rasmussen, M. *et al.* Ancient human genome sequence of an extinct Palaeo-Eskimo. *Nature* **463,** 757–762 (2010).
18. Tackney, J. C. *et al.* Two contemporaneous mitogenomes from terminal Pleistocene burials in eastern Beringia. *Proc. Natl. Acad. Sci.* 201511903 (2015). doi:10.1073/pnas.1511903112
19. Verdu, P. *et al.* Patterns of Admixture and Population Structure in Native Populations of Northwest North America. *PLoS Genet.* **10,** e1004530 (2014).
20. Patterson, N. *et al.* Ancient Admixture in Human History. *Genetics* **192,** 1065–1093 (2012).
21. Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19,** 1655–1664 (2009).
22. Skotte, L., Korneliussen, T. S. & Albrechtsen, A. Estimating Individual Admixture Proportions from Next Generation Sequencing Data. *Genetics* **195,** 693–702 (2013).
23. Malaspinas, A.-S. *et al.* bammds: a tool for assessing the ancestry of low-depth whole-genome data using multidimensional scaling (MDS). *Bioinformatics* **30,** 2962–2964 (2014).
24. Pickrell, J. K. & Pritchard, J. K. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genet.* **8,** e1002967 (2012).
25. Dyke, A. S., Moore, A. & Robertson, L. Deglaciation of North America. (2003).
26. Pedersen, M. W. *et al.* Postglacial viability and colonization in North America's ice-free corridor. *Nature* (2016). doi:10.1038/nature19085
27. Kari, J. M. & Potter, B. A. *The Dene-Yeniseian connection*. (University of

Alaska Department of Anthropology/Alaska Native Language Center, 2011).
28. Steinrücken, M., Kamm, J. A. & Song, Y. S. Inference of complex population histories using whole-genome sequences from multiple populations. *bioRxiv* (2015). doi:10.1101/093468
29. Kamm, J. A., Terhorst, J. & Song, Y. S. Efficient computation of the joint sample frequency spectra for multiple populations. *J. Comput. Graph. Stat.* **26,** 182–194 (2016).
30. Goebel, T. The 'microblade adaptation' and recolonization of Siberia during the late Upper Pleistocene. *Archeol. Pap. Am. Anthropol. Assoc.* **12,** 117–131 (2002).

Figure 1. Genetic affinities between USR1, present-day Native Americans, and world-wide populations. a. *f3* statistics of the form *f3(San; X, USR1),* for each population in the genotype panel. Warmer colors represent greater shared drift between a population and USR1. **b.** *D*-statistics of the form *D(Native American, Aymara; USR1, Yoruba)* (points). The Andean Aymara were used to represent SNA. *: Native American populations with Asian admixture (|*Z*| for *D(H1, Aymara; Han, Yoruba)>3.3*) (Figure S5a). Error bars represent 1 and ~*3.3* standard errors (*p*-value~0.001). Native American populations were grouped by language family [1]. **c.** Quantile-quantile plot comparing observed *Z*-scores to the expected normal distribution under the null hypothesis (H_0), for all possible *D(Nat. Am., USR1; Siberian1, Siberian2)*. Colors correspond to the Z-score obtained for *D(H1, Aymara; Han, Yoruba)*. The expected normal distribution under the null hypothesis was computed for all groups jointly (SI Section 10.4). Thick and thin lines represent a *Z*-score of ~*3.3* (*p-val~0.001*) and a *Z*-score of ~4.91 (*p-val~0.01* after applying a Bonferroni correction for 11,322 tests). The bottom-right panel shows the expected tree under the null hypothesis. **d.** Admixture proportions estimated by *ADMIXTURE* [37] assuming *K=20* ancestral populations. Bars represent individuals, and colors represent admixture proportions from each ancestral component. Admixture proportions in ancient genomes (wider bars) were estimated using a genotype likelihood-based approach [38].

Figure 2. Possible geographic locations for the USR1 and NNA-SNA splits. We propose two possible locations for the split between USR1 and other Native Americans: the Old World (**A**, **C**, **E**) and Beringia (**B**, **D**); and three possible locations for the NNA_SNA split: the Old World (**E**), Beringia (**C**, **D**), and North America south of Beringia (**A**, **B**). Schematics show estimated glacial extent ~14.8 kya. Dashed lines represent the Native American migration south of eastern Beringia, but they do not correspond to a specific migration route Model discussion (SI 20) is based on extant archaeological evidence and inferred demographic parameters: a USR1-NNA+SNA split ~20 kya with ensuing moderate gene flow and a NNA-SNA split ~15 kya (SI 18-19)..

Figure 3. A model for the formation of the different Native American populations. We fitted an admixture graph by sequentially adding admixed leaves to a 'seed' graph including the Yoruba, Han, Mal'ta, Ket, USR1, Anzick1 and Aymara genomes. For each 'non-seed' admixed group, we found the pair of edges that produced the best-fitting graph,

based on the fitting and maximum |*Z*| scores (3.27 for this graph). Ellipse-shaped nodes: sampled populations; box-shaped nodes: metapopulations; *: single high-depth ancient genome. **: single low-depth genome. †: subgraphs whose structure we were unable to resolve due to sequencing and genotyping error in the Saqqaq genome (SI 17). Sample sizes and locations are shown at the top.

Figure 4. USR1 demographic history in the context of East Asians, Siberians and other Native Americans. a. *SMC++* inferred effective population sizes with respect to time for Athabascans (NNA), Karitiana (SNA), Han, Koryaks and USR1 (SI 19.1). We used these demographic histories as a basis for fitting a joint model for these populations. **b.** A 'backbone demography' was fitted excluding USR1 using *momi2*, an SFS-based maximum likelihood approach (Figure S27), along with the most likely join-on point for USR1 onto the backbone demography (SI 19). We show the likelihood heatmap for the latter; warmer colors correspond to a higher likelihood of USR1 joining at a given point. These estimates agree with those obtained through *diCal2*, a method based on haplotype data (SI 18).